Measurement of the Top Mass in the Lepton + Jets Channel Moriond 2004: The Low Bias Template Method

K.M. Black, M. Narain

March 22, 2004

Abstract

A measurement of the top quark mass in the lepton+jets channel using Run II data is presented. The data sets considered correspond to integrated luminosities of 168.7 pb^{-1} and 158.4 pb^{-1}, in the e+jets and mu+jets channels, respectively. After selection, a two-constraint kinematic fit to the ttbar hypothesis is performed and an event discriminant between signal and background is built from four topological variables using a template based likelihood method. We measure the top mass to be 168 $\pm$ 13 GeV (stat) $\pm$ 8 GeV (sys)

1   Introduction

In the standard model the top quark is predicted to decay into a W boson and a bottom quark with a branching ratio of approximately 1. The W can decay hadronically into two quarks or leptonically into a charged lepton and a neutrino. In this note, we consider $t\bar{t}$ events in which one W decays hadronically and the other leptonically: the "lepton + jets" channel. Further, given that τ s are difficult to identify experimentally, we restrict ourselves to the electron + jets and muon + jets channel.

This note describes one of the methods to extract the top mass from our data sample. The method is based upon the Run I method [4] which uses a kinematic fit of the event four-vectors. Then from the Monte Carlo, templates are constructed with various hypothesized top

masses. As well a template constructed from the main background, W events with multiple jets from initial state radiation is produced. The top mass is then extracted from the data using a likelihood fit which estimates the most likely value of the top mass and the number of $t\bar{t}$ events in the sample.

2   Event Selection

For the purpose of this analysis, we subdivide the lepton+jets final state into e+jets and μ+jets. The current analysis is purely "untagged" and it does not consider yet any b-tagging information (neither soft-lepton nor lifetime tagging).

The data sample considered is preselected following closely the criteria used in the measurement of the $t\bar{t}$ production cross-section [3], with the exception that no veto is applied on events with soft-muons. This cut is motivated to make orthogonal data sets for the cross-section analysis and using this cut reduces the signal to background ratio of the sample. In this document, we only summarize the event preselection. For a detailed description of the efficiency and motivation element of the event selection please see the top production note for the winter conferences [3].

2.1   Object Identification

In this section we describe the basic physics object definitions.

Primary Vertex Selection: In order to ensure high quality of the reconstructed vertex it is required that the primary vertex be within the SMT fiducial region ($|PV_z| < 60$ cm) with a minimum of track multiplicity of 3.

Electrons: The identification of electrons follows from the initial EM cluster reconstruction and also adds the requirement that:

1. $f_{EM} > 0.9$.

2. $f_{iso} < 0.15$.

3. hmx8 $\chi^2 < 75$.

4. the electron has a track match from the central tracker.

5. the electron passes a likelihood cut of $L_{CC} > 0.75$ [2].

6. the electron has $p_T > 15$ GeV.

7. the electron is in the central calorimeter with detector $|\eta| < 1.1$.

Muons: Muon reconstruction is described in section VI of [1] Further we require that:

1. We selected medium muons that have a track match associated and $N_{seg} = 3$.

2. The muon satisfies the cosmic ray veto.

3. The muon is well isolated from any jet $\Delta R(\mu, jet) > 0.5$.

4. The muon pass a calorimeter halo isolation described in the muon [1].

5. The muon is consistent with originating from the primary vertex in both the transverse plane and in the z direction: $\frac{|dca|}{\sigma_{dca}} < 3$ and $\Delta Z(\mu, PV) < 1$ cm.

Jets: Jets are reconstructed according to the Run II cone algorithm with $\Delta R = 0.5$ using the new T42 algorithm. To suppress isolated electromagnetic particle contributions and instrumental problems we require that:

1. $0.05 < f_{EM} < 0.95$.

2. $f_{CH} < 0.4$.

3. hot cell fraction < 10.0.

4. n90 > 1.

$\not{E}_T$: We follow the general top group procedure for defining the transverse missing energy in an event found in [1].

2.2 Preselction

The event selection can be divided in to two steps: preselection and final event selection. As stated before the preselction follows closely the event preselection for the topolgocial cross section measurement for Moriond 2004 [3]. As mentioned earlier, we do not reject events with a soft muon. Soft muons can originate from the decay of b quarks in the $t\bar{t}$ decay, so keeping these events enriches our sample with $t\bar{t}$ events.

Muon + Jets Preselection: The events first must pass either one of the two signal triggers MUJT20l2M0 and MUJT25L2M0 [3] For trigger versions 8-11 the trigger used is MUJT20L2MO, for trigger version 12 MUJT25L2M0. This requires an isolated muon and one jet

3

with $P_T > 20$ or GeV $P_T > 25$ GeV. The offline selection requires the events to have:

1. An isolated medium muon with $P_T > 20$ GeV.

2. 4 jets with $P_T > 15$ GeV and $|\eta| < 2.5$.

3. $\not{E}_T > 17$ GeV.

4. $\Delta\phi(\mu, \not{E}_T) > 1.2$ - $\not{E}_T \times 1.2/38$

5. $\Delta\phi(\mu, \not{E}_T) < 1.3 + \not{E}_T \times (\pi - 1.3)/24$

6. $\Delta\phi(leading jet, \not{E}_T) < 2.2 + \not{E}_T \times (\pi - 2.2)/26$

The triangular $\Delta\phi$ cuts are designed to reduce the multi-jet fake rate while keeping the efficiency of the $t\bar{t}$ as high as possible.

For trigger versions 8-11 the trigger used to select electron + jet evetns was EM152J15, for trigger version 12 E1SHT152J2 [3].

Electron + Jets Preselection:

1. An electron which passes all quality cuts with $P_T > 20$ GeV.

2. 4 jets with $P_T > 15$ GeV and $|\eta| < 2.5$.

3. $\not{E}_T > 20$ GeV.

4. $\Delta\phi(\not{E}_T, e) > 1.7$ - $\frac{1.7 \times \not{E}_T}{80 GeV}$.

The multi-jet background is caused primarily by events the lepton is misidentified and the missing transverse momentum of the event is signficantly mismeasured. In order to obtain a sample that is rich in these type of misidentified results reverse the quality selection on the charged lepton but keep the rest of the preselection the same. In the case of the cross-section analyis [3] they reverse the selection on the missing transverse momentum as well. However, the missing transverse momentum is an important quantity for the fit mass of the event so that reversing this selection would result events which do not well reproduce our background in this crutial aspect .In the electron + jets case, one jet from a multi-jet event can have a particularly high electromagnetic fraction and pass the lepton quality selection. In the muon + jets case a charged pion or other hadron from a hadronic jet can decay producing a muon and be misidentifed as a muon from the hard scattering. In order to isolate events that are high in multi-jet background, the same preselection was taken except that in the electron case the electron was required to fail the electron likelihood and in the muon case the muon is required to fail the isolation criterion.

The efficiency for $t\bar{t}$ events to pass these reverse cuts is less then a few percent and therefore we can conclude that given the relative cross-section for multi-jet and $t\bar{t}$ events this reverse lepton selection produces a sample very rich in background.

In contrast to the Matrix Element method, we use events which have 4 or more jets. This increases the statistics in the sample, but currently we only consider the leading 4 jets as coming from the top in the mass fitting. Other jets in the events are treated as if they came from initial state radiation and are not included in the final mass fitting.

Two other cuts are used to increase the signal to background ratio which are not used in the current cross-section analysis. In the electron + jets channel, the fit estimated multi-jet background is significantly larger then in the muon + jets channel (20 % of the sample as opposed to 7 % of the sample). A cut which was used in Run I whose motivation was to remove qcd background is a cut on the "$E_T W \equiv |P_T \ electron| + \not{E}_T$". The second cut, which is used for both channels and was use in the Run I mass analysis [4] is based on a variable called ht2. This variable is simply the sum of the jet pt for all jets with $|\eta| < 2.5$ and Jet $P_T > 15$ GeV excluding the leading jet. Figure 1 shows the distribution from Monte Carlo for $E_T W$ while figure 2 shows the distribution of ht2. For this analysis we further require that:

1. For the electron + jet channel events have $E_T W > 65$ GeV.

2. For both channels we require that ht2 > 90 GeV.

3. For both channels we require that the lowest χ^2 permutation has converged and has a good fit with $\chi^2 < 10$.

In order to cut on Ht2 ,we need to undertand how well the variable is modeled with MC. Since the W + 4 jet statistics are low and the data will have a top contribution, we first take a look at the W + 3 jet data. We require the full event selection except for the ht2 requirement, and the χ^2 of the fit (since these are 3 jet events). However, if we were to just look at these events we know that there would be two major contributions: W's produced along with 3 jets from initial state radiation, and multi-jet events where the the there is misidentification of the lepton (a jet passes the electron criterion in the electron + jets case or a muon comming from the decay of a hadron inside a jet passes the isolation criterion by decaying at a large angle relative to the jet direction).
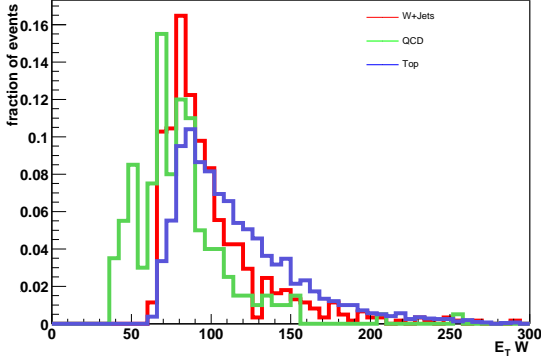
Figure 1: E_T W of top and W+ jets Monte Carlo compared to the qcd sample

Therefore, in order to extract the W + 3 jet contribution from the data we use the first matrix method described in [3]. In this case we divide the sample in to a loose and a tight sample based on the quality of the lepton. In the electron + jets channel we use the electron likelihood and in the muon +jets channel we use the muon isolation variable. If the event passes the tight lepton requirement it becomes part of the tight sample.

The matrix method is well known, but to sumarize if we have two samples with two contributions we write that the loose and the tight sample are composed as:

$$N_l = N^W + N^{QCD} \tag{1}$$

$$N_t = \epsilon_W N^W + \epsilon_{QCD} N^{QCD} \tag{2}$$

where N_l N_t are the number of events in the loose and tight samples, respecitvely and N^W and N^{QCD} are the number of events from W + jets proccess and multijet proccesses, respectively. Here, ϵ_W is the efficiency of W + 3 jet events to pass the tight requirement and ϵ_{QCD} is the efficiency of the multijet events to pass the tight requirement. If we measure these from data and Monte Carlo then we can solve for N^W and N^{QCD}. We take ϵ_W and ϵ_{QCD} from [3].

$$N_W = \frac{N_t - \epsilon_{QCD} N_l}{\epsilon_W - \epsilon_{QCD}} \tag{3}$$

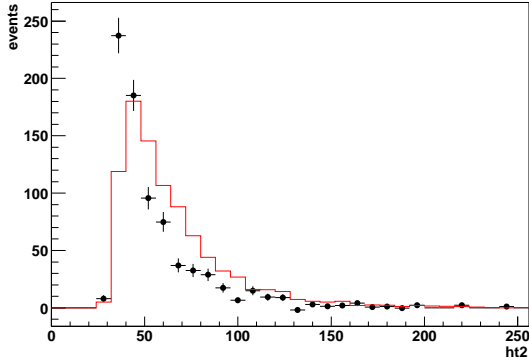$$N_{QCD} = \frac{\epsilon_W N_l - N_t}{\epsilon_W - \epsilon_{QCD}} \tag{4}$$

6

Figure 2: Ht2 distribution for the W+3 jet bin exclusively in the electron + jets channel. The matrix method is used to calculate the contribution from W+3 jets in data while the red line represents the monte carlo prediction.

Then on a bin by bin basis from the tight and loose sample we compute the number of W + 3 jet events and compare with with the Monte Carlo prediction. The results for the electron channel are shown in figure 2. In figure 3 the results from the muon channel are shown.

There are two things that are obvious from these distributions. First, both show that the Monte Carlo is has a systematically harder jet p_T spectrum.

In order to address this point two quick invesigations were made. The first study was to increase the jet P_T requirement to 25 GeV. The results are shown inf figures 4 and 5.

Of course the statistical fluctuation for both data and monte carlo has increased dramatically. However, there is a trend towards better agreement. As well, W + jets with lower parton level cuts, a different hard scattering scale, and the correct treatment of multiple parton interactions was generated. Unfortunately, at this time, only W + 2 jet and W + 4 jet monte carlo is available. Using the W + 4 jet monte carlo with events that happen to only have 3 jets due to acceptance and reconstruction would obviously bias the monte carlo to a harder P_T spectrum. Instead, we compare the W + 2 jet monte carlo with events that happen to have 3 jets due to radiaton.

The results are shown in figures 6 and 7.

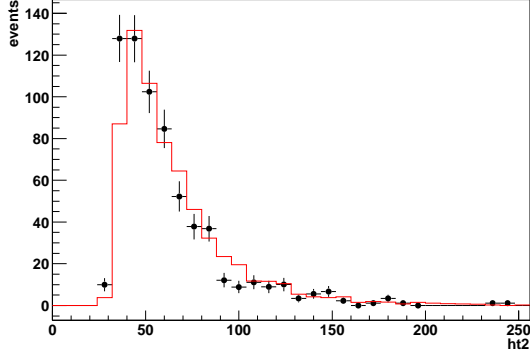Clearly, the monte carlo that is generated with the lower parton

Figure 3: Ht2 distribution for the W+3 jet bin exclusively in the muon + jets channel. The matrix method is used to calculate the contribution from W+3 jets in data while the red line represents the monte carlo prediction.
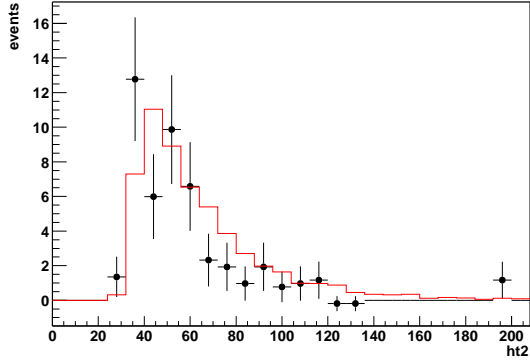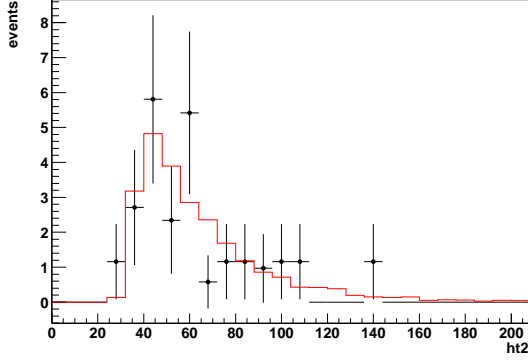

Figure 4: Ht2 distribution for the W+3 jet bin (exclusively) in the electron + jets channel. The matrix method is used to calculate the contribution from W+3 jets in data while the red line represents the monte carlo prediction. In these plots, every jet is required to have at $P_T > 25$ GeV.
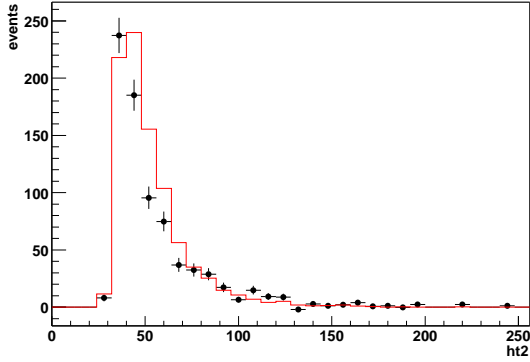
Figure 5: Ht2 distribution for the W+3 jet bin (exclusively) in the muon + jets channel. The matrix method is used to calculate the contribution from W+3 jets in data while the red line represents the monte carlo prediction. In these plots, every jet is required to have at $P_T > 25$ GeV.

Figure 6: Ht2 distribution for the W+3 jet bin exclusively in the electron + jets channel. The matrix method is used to calculate the contribution from W+3 jets in data while the red line represents the monte carlo prediction with lower parton level cuts, tune A, and a different hard scattering scale. Note that this was generated to be a W + 2 jet sample.
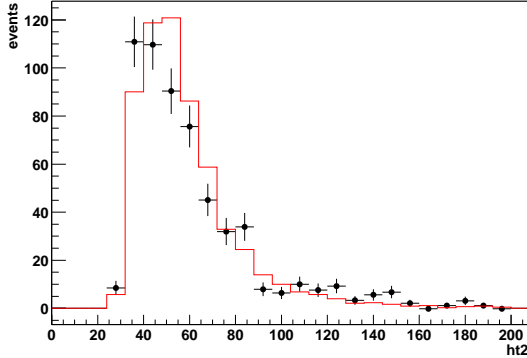
9

Figure 7: Ht2 distribution for the W+3 jet bin exclusively in the muon + jets channel. The matrix method is used to calculate the contribution from W+3 jets in data while the red line represents the monte carlo prediction with lower parton level cuts, tune A, and a different hard scattering scale. Note that this was generated to be a W + 2 jet sample.

level cuts, parton level interactions, and a smaller hard scattering scale gives better agreement. However, simply put, the agreement is obviously not satisfactory. Given the relatively (though much lower statistic) agreemet with the higher jet P_T requirement in conjunction with the other studies displayed in this note it appears that the description of lower pt jets (< 20 GeV) is at least partially to blame.

The efficiency for the $E_T W$ selection is:

1. 95.7 $\pm$ 0.3 % for $t\bar{t}$.

2. 92.2 $\pm$ 0.5 % for W +jets.

3. 48.9 $\pm$ 6.0 % for multijet events.

The efficiency for the ht2 selection is:

1. 97.3 $\pm$ 0.3 % for $t\bar{t}$.

2. 50.1 $\pm$ 0.4 % for W +jets.

3. 54.2 $\pm$ 6.2 % for qcd.

The Ht2 distributions seems to invite a higher threshold for this selection. However, as can be seen from figure 9, the correlation between fitmass and Ht2 is quite high. Therefore in order to avoid biasing the fit mass distribution and as discussed in [4] ultimately leading to a
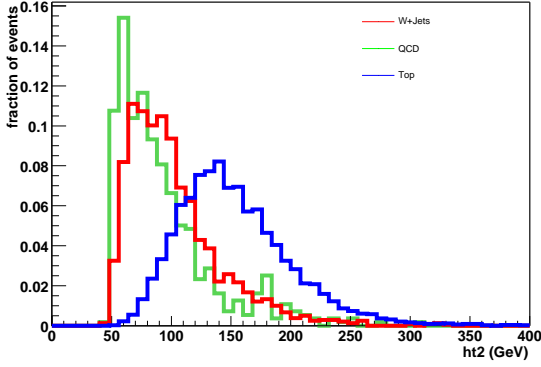
Figure 8: Ht2 of top and W + jets Monte Carlo compared to the qcd sample.

reduced sensitivity of the top mass this selection is intentionally not optimized for a maximum signal to background. For a true top quark mass greater then 150 GeV, the ht2 > 90 GeV at most removes about 12 % of the top signal while removing approximately half of the background.

As described in the next section, the Ht2 selection removes about half of the electron + jets canidates and about 40% of the muon + jet sample. Therefore, it is important that we take a close look at the events that we are discarding from the analysis in order to make sure that we are not throwing away top events. In order to do so, we examined several variables and compared them to the predicted shape and normalization from the expected sample composition. In figures 10-13 we compare the leading jet p_T, the isolated lepton p_T, the missing transverse momentum, and the fit mass of the events. All seem to be well modeled by the background hypothesis, and clearly show no evidence for top as expected from the selection efficiencies.

The conclusion that the top fraction in the sample with Ht2 < 90 GeV is also supported by the small number of events that have a jet identified as comming from a b quark. There are 5 events which have one jet that is identified out as a b-jet. Two of these events would fail the low bias discriminant selection described later. The fit mass of these events is shown in figure 14.
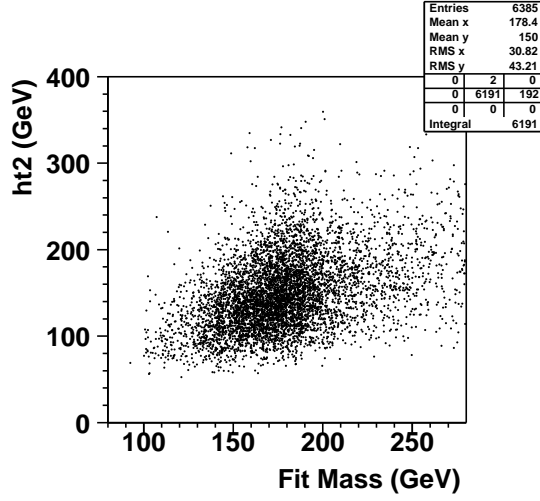
Figure 9: Scatter plot of Ht2 (y axis) and fit mass (x-axis) for a $t\bar{t}$ sample generated with 175 GeV top mass. Obviously the correlation is quite large.
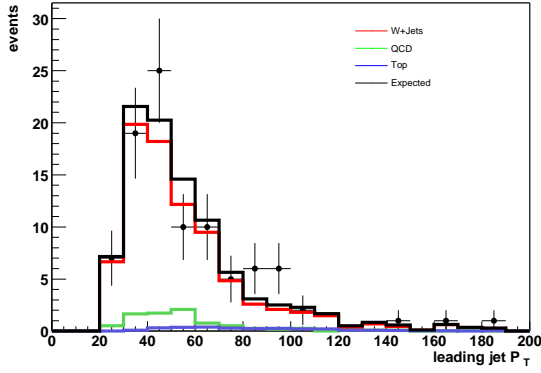

Figure 10: The leading jet p_T of the events failing the Ht2 $>$ 90 GeV selection compared to the expected sample composition.
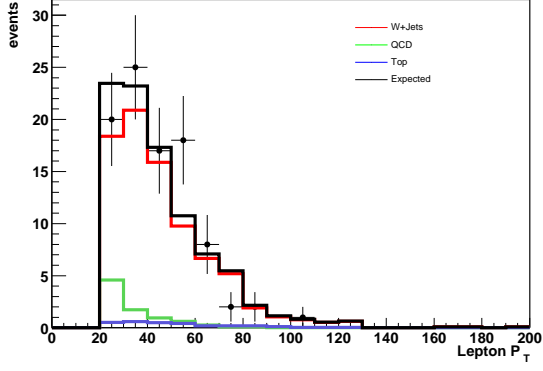
Figure 11: The isolated lepton p_T of the events failing the Ht2 > 90 GeV selection compared to the expected sample composition.
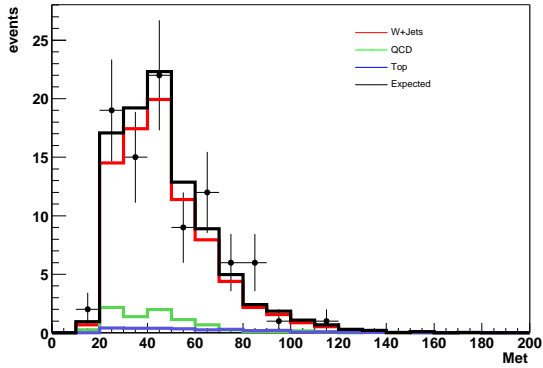

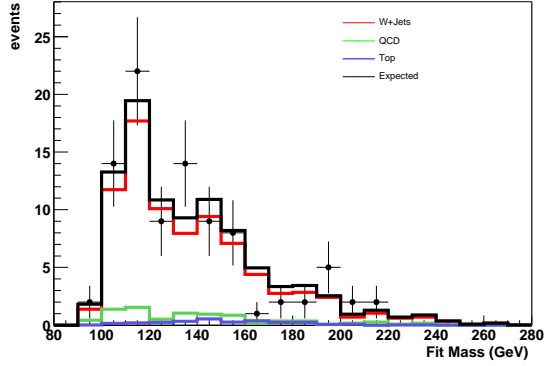Figure 12: The $\not{E}_T$ of the events failing the Ht2 > 90 GeV selection compared to the expected sample composition.

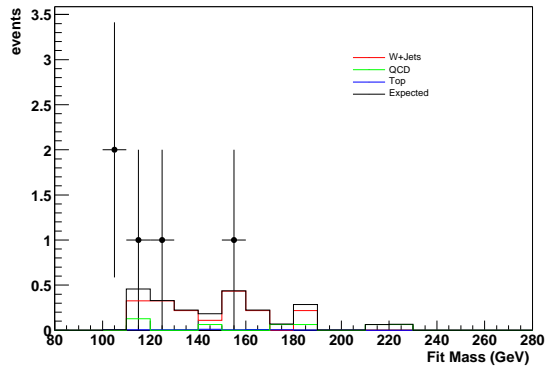Figure 13: The fit mass of the events failing the Ht2 > 90 GeV selection compared to the expected sample composition.

Figure 14: The fit mass of the tagged events failing the Ht2 > 90 GeV selection compared to the expected sample composition.

2.3   Event Yield

The starting point for the event selection is the topological cross-section. These analysis were completed with 143.9 pb^{-1} and 141.2 pb^{-1} of integrated luminosity for the μ +jets and e + jets channels, respectively. However, since that time period about 15 pb^{-1} have been added.

After the $E_T W$ selection in the electron + jets channel we find 126 events. S. Kulik then uses the two matrix methods described in [3] to estimate that the electron + jets. channel has the following composition:

1. 38.4 ± 8.1 $t\bar{t}$ events.

2. 76.2 ± 3.0 W+jet events.

3. 11.0 ± 1.6 multi-jet events where one jet passes the electron identification.

In the muon + jets channel we estimate that there are 104 events with the estimated composition shown below.

1. 28.0 ± 7.1 $t\bar{t}$ events.

2. 67.6 ± 2.6 W+jet events.

3. 8.2 ± 1,0 multi-jet events where a muon from one jet passes the isolated muon identification.

After requiring that the the events also have Ht2 > 90 GeV we find 65 events in the μ + jets channel. Given the efficiencies for this requirement shown above we can then estimate that the sample has the following composition:

1. 27.3 ± 6.8 $t\bar{t}$ events.

2. 33.8 ± 1.3 W+jet events.

3. 4.4 ± 0.5 multi-jet events where a muon from one jet passes the isolated muon identification.

In the electron + jets channel we find 72 events after the ht2 > 90 GeV requirement. Again from the efficiencies above we can estimate the sample composition:

1. 37.6 ± 7.8 $t\bar{t}$ events.

2. 38.1 ± 1.5 W+jet events.

3. 6.0 ± 0.9 multi-jet events where one jet passes the electron identification.

After all preselection cuts the signal to background ratio is estimated at approximately 1:1.

3 Topological Discrimination

There are two dominant backgrounds in the $t\bar{t}$ lepton + jets channel. The first is W events which are accompanied by multiple jets. These jets can come from the hard scattering event, initial or final state radiation, or the underlying event. The other major background comes from QCD multi-jet events where one jet is mis-identified as an electron and the $\not{E}_T$ is significantly mis-measured. The π^0 content of a hadronic jet can be high causing a narrow, highly electromagnetic deposition of energy in the calorimeter. As well a π^0 can be mistakenly associated with a charged particle track. Either of these can cause a jet to be mis-identified as an electron.

Since the top is such a heavy object, variables such as the scalar sum of the jet p_T can give quite good discrimiantion between $t\bar{t}$ and background. By making a selection on event topology, it was seen that the resulting events had reconstructed top masses that were highly biased. As well the shape of the irreducible background was indistinguishable from the signal making it difficult to separate the two and remove the bias. The approach that was taken makes use of topological variables describing the shape of the event to form a discriminant between signal and background. Rather than making an event selection on this discriminant, this information was incorporated with the fit mass in the final likelihood procedure to extract the top quark mass [4] [5].

Since there was significant effort in finding appropriate variables that allow good discrimination without the introduction of significant bias the Run I procedure was adopted as the starting point in this analysis.

3.1 Variables

In Run II operating parameters are different from Run I. Most notably for $t\bar{t}$ production, the center of mass energy is higher (from 1.8 to 1.96 TeV) which results in the $t\bar{t}$ cross-section to be 40% higher. Therefore,

one might not expect the same variables that were optimal in Run I to continue to be so in Run II. Nonetheless, these variables are taken as the starting point of this analysis. As it matures, other variables and approaches will be developed.

The four topological variables considered here are:

1. The reconstructed $\not{E}_T$.

2. $\mathcal{A} \equiv \frac{3}{2} \times$ smallest eigenvalue of $\mathcal{P}$

3. $H'_{T2} \equiv \frac{H_{T2}}{H_\parallel}$

4. $K'_{Tmin} \equiv \frac{(min \ of \ 6 \ \Delta R_{jj}) \cdot E_T^{lesser \ j}}{E_W^T}$

These definitions need some explanation. First, the term reconstructed $\not{E}_T$ is used for the following reason. Because of mis-measurement of the lepton momentum or the $\not{E}_T$ the transverse mass can exceed the W pole mass M_W. This could also result naturally since the W line-shape has a Breit-Wigner tail. This is seen to happen approximately 20% of the time. As described in the lepton + jets kinematic fitting section, the p_z component of the neutrino is found by forcing the invariant mass of the charged lepton and $\not{E}_T$ to M_W . In this case, allowing the transverse mass to exceed the W mass is impossible. Attempting to do so results in only imaginary solutions for p_z^ν which obviously cannot be used to fit the event. In order to avoid this, the $\not{E}_T$ is re-scaled such that $M_T = M_W$ [5] . Eventually, this procedure needs to be rethought. Since the distribution of W mass is a Breit-Wigner lineshape, a future approach would be to produce an algorithm which finds the most probable value of the W mass for that event drawn from the Breit-Wigner [5] .

$\mathcal{A}$ is the aplanarity of the event. $\mathcal{P}$ is the normalized momentum tensor of the event derived from the momenta of the jets and the reconstructed W. It is defined by:

$$\mathcal{P}_{ij} \equiv \frac{\sum_a p_{a,i} p_{a,j}}{\sum_a |\vec{p}_a|^2} \tag{5}$$

where i and j label the **spatial** components of the momentum, and a runs over all jets and the reconstructed W. $\mathcal{A}$ is defined as $\frac{3}{2}$ the smallest eigenvalue of $\mathcal{P}$ and has a range of 0 and 0.5. It can be shown [6] that events that have highly aplanar momentum vectors have large values of aplanarity.

$H_\parallel$ is the scalar $|p_z|$ of the jets, isolated lepton, and the neutrino. H_{T2} is the sum of the $|p_T|$ of the jets excluding the leading jet. H'_{T2} is

17

seen to give rather good discrimination as well as a small correlation with the fit mass. Although variables such as H_T have notably better discrimination power, it was the Run I experience that the correlation with the top mass was unacceptably large [5].

K'_{Tmin} is a measure of the jet separation folded together with "transverse energy" of the reconstructed W. ΔR_{ij} is the distance between jet i and jet j in $\eta - \phi$ space. Of the six possible ΔR_{ij} between the four leading jets, the smallest is chosen. $E_T^{lesserj}$ is the smaller of the two jets E_T 's. Note that the transverse energy of the W is defined as the sum of $\left|p_T^l\right|$ and $|p_T^\nu|$.

Note that two other variables was seriously considered during Run I: $\eta_{rms} = \sqrt{\langle p_T\ weighted\ \eta^2 \rangle}$ and sphericity. Both variables as well were seen to give good discrimination power, however both had a high correlation with the other variables used (most notably H'_{T2} and $\mathcal{A}$) and hence did not contribute very much to the overall discrimination power. When the studies were repeated in Run II, identical results were found. Note that the cross-section analyis for Moriond 2004 uses sphericity in place of $\not{E}_T$. Comparing likelihood distributions with [3] gives similar discrimination from background.

3.2   Distributions

In order to understand the topology of both signal and background, we examine the four selected topological variables.

Figure 15 shows the distribution of $\mathcal{A}$ obtained for a background sample of W+jets and $t\bar{t}$ Monte Carlo. Since after the E_T W selection the fraction of multijet events is quite in the background 10 % the discriminant is built only using W + jet monte carlo to maximimize the discrimination against the major background.

Figure 16 shows the distribution for H'_{T2} . As discussed previously, the H_T of the event has a extremely high correlation with the fit mass. In [5] notes that even H_{T2} has a correlation of approximately 0.5 with the fit mass (a number with which we are consistent). This is fairly straightforward to understand as higher p_T jets yield higher fit masses. By dividing this by another variable, $H_\parallel$ that is also correlated with the mass one can reduce this correlation.

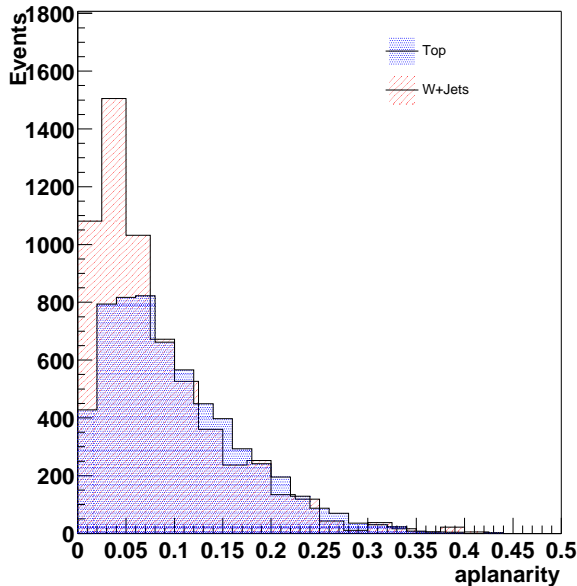In figure 17, the distribution for K'_{Tmin} is shown.

Figure 15: Aplanarity for top and w+jet MC.

3.3 Correlations

There are two types of correlations which are important in constructing a low bias likelihood. The first, and most important is the correlation of the variables that form the likelihood and the fitmass for the events. The purpose of constructing the low bias likelihood is to determine a way of separating the signal from background without either loosing sensitivity or biasing the measurement in an unknown way. Figures 19 through 22 show the correlation of the four topological variables with the fit mass.

The correlations between the topological variables and the fitmass is summarized here (for the top sample):

1. $\not{E}_T$: 0.06

2. $\mathcal{A}$: -0.09

3. H'_{T2}: 0.02

4. K'_{Tmin}: 0.03

It is important to check that the variables are not significantly correlated with each other nor the fit mass. In order to do so we compute
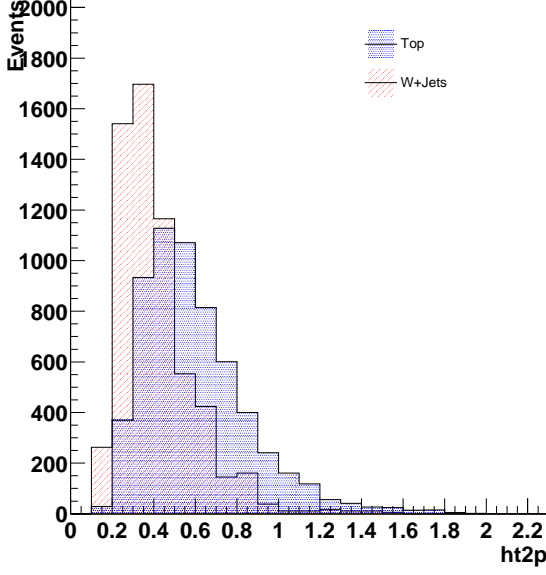
Figure 16: H'_{T2} distributions for $t\bar{t}$ and background.

the standard Pearson product linear moments. For the correlations among themselves, this is best represented by a symmetric matrix. The elements are calculated according to:

$$C_{ij} = \frac{\langle (L_i - \bar{L}_i)(L_j - \bar{L}_j) \rangle}{\sigma_i \sigma_j} \qquad (6)$$

if i = (1, 2, 3, 4) = ($\not{E}_T$, $\mathcal{A}$, H'_{T2}, K'_{Tmin}) - we obtain:

$$\mathcal{C} = \begin{pmatrix} 1 & -0.05 & 0.03 & -0.04 \\ -0.05 & 1 & 0.37 & 0.12 \\ 0.03 & 0.37 & 1 & 0.17 \\ -0.04 & 0.12 & 0.17 & 1 \end{pmatrix} \qquad (7)$$

The only variables with significant correlation are $\mathcal{A}$ and H'_{T2} . The correlations between the various variables are displayed in figures 23-29.
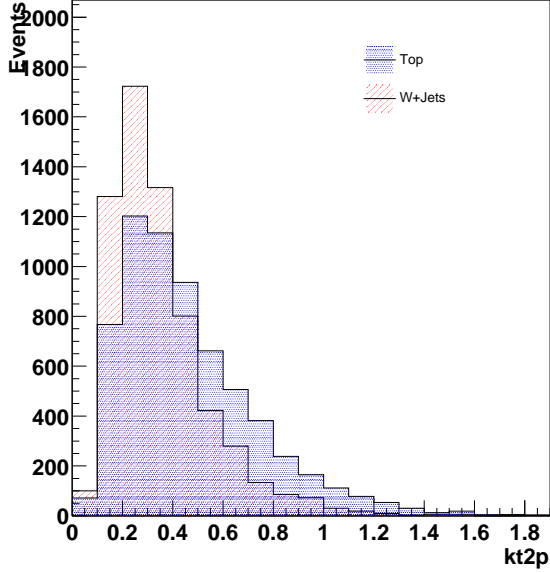
Figure 17: K'_{Tmin} distributions for $t\bar{t}$ and background.

3.4 Forming the Discriminant

Now that we have shown that we can obtain discrimination between our signal events and background we must develop a procedure to obtain maximal discrimination by combining these variables into a likelihood. We follow the work described in [5]. For each event we compute a discriminant which is mapped between 0 and 1 from the topological variables. The closer to 1, the more top like. This procedure is described here and is a bit intricate.

First, the topological variables are not used directly. A set of functions is used to map the variables such that the distributions are more smoothly varying. The motivation for this step is that when performing fits to the topological variables the fits are highly effected by the large variations. In particular, statistical fluctuations in rapidly varying regions can cause bad fits and poor understanding of the data. The set of functions that is used are (again following[5]):

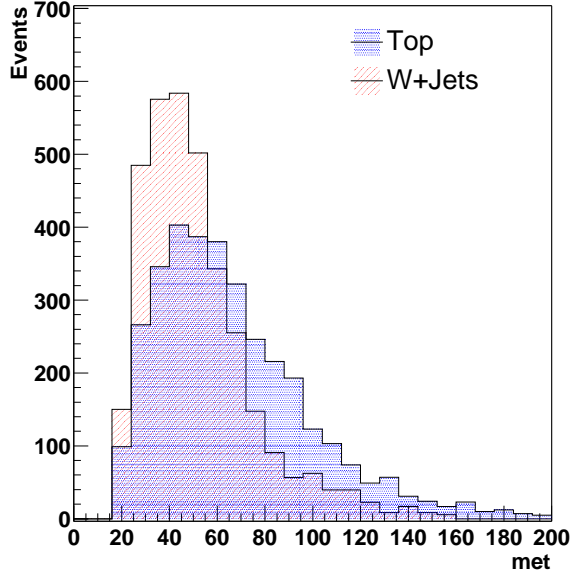1. $\mathcal{A}$: $e^{-11\mathcal{A}}$

2. H'_{T2}: $\ln H'_{T2}$

21
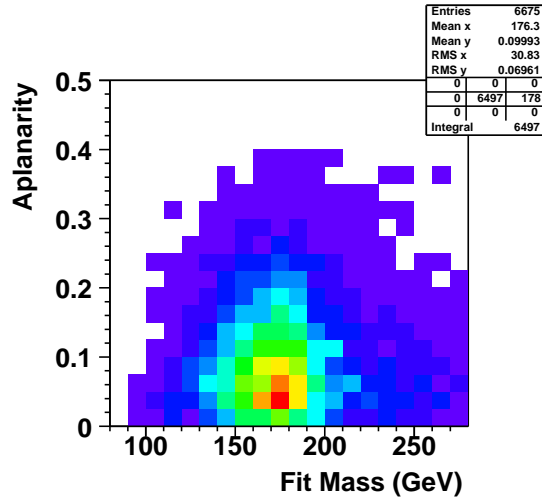
Figure 18: $\not{E}_T$ distributions for $t\bar{t}$ and background.

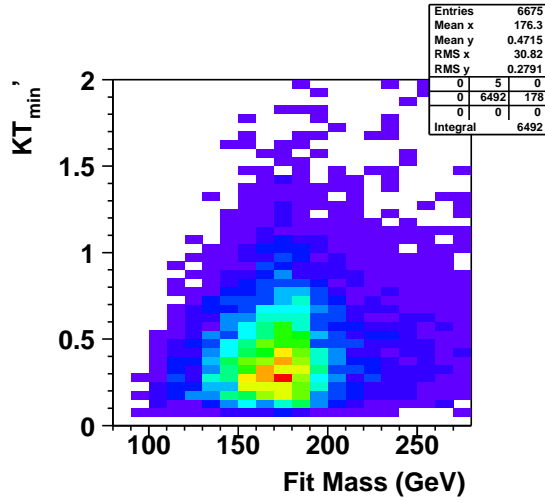Figure 19: Correlation between the fitmass and $\mathcal{A}$.

Entries		6675
Mean x		176.3
Mean y		0.4715
RMS x		30.82
RMS y		0.2791
0	5	0
0	6492	178
0	0	0
Integral		6492

Figure 20: Correlation between the fitmass and K'_{Tmin}.

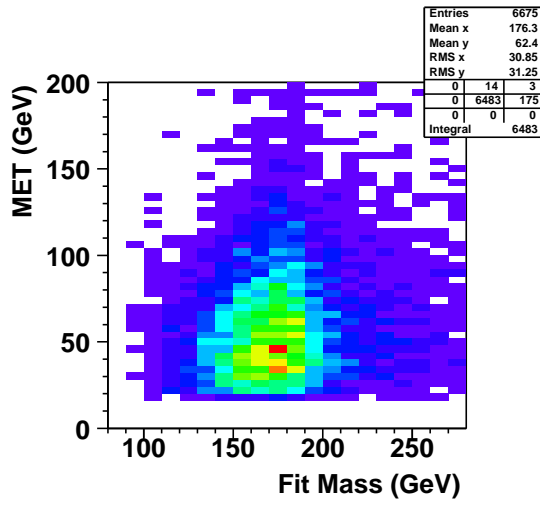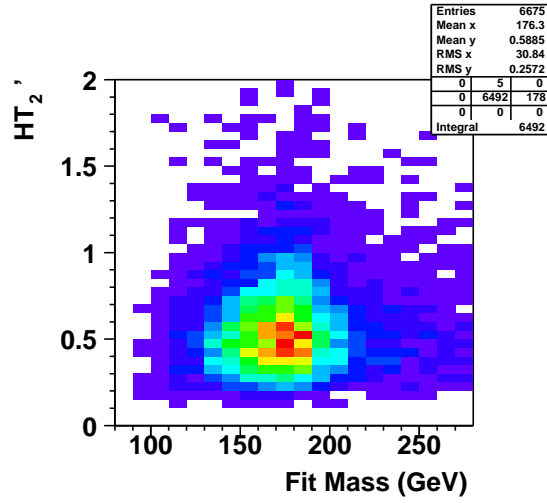|		
Entries		6675
Mean x		176.3
Mean y		62.4
RMS x		30.85
RMS y		31.25
0	14	3
0	6483	175
0	0	0
Integral		6483

Figure 21: Correlation between the fitmass and $\not{E}_T$.

Figure 22: Correlation between the fitmass and H'_{T2}.
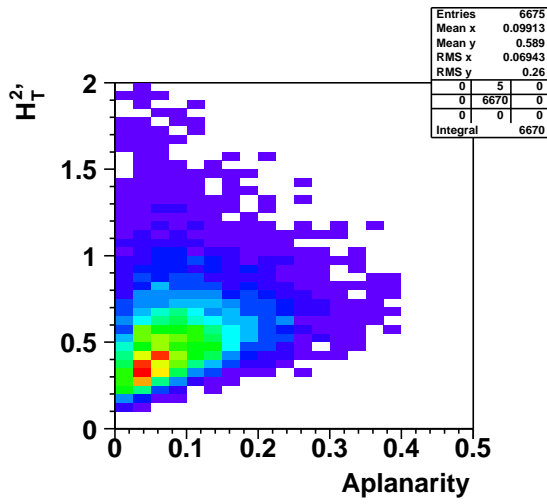
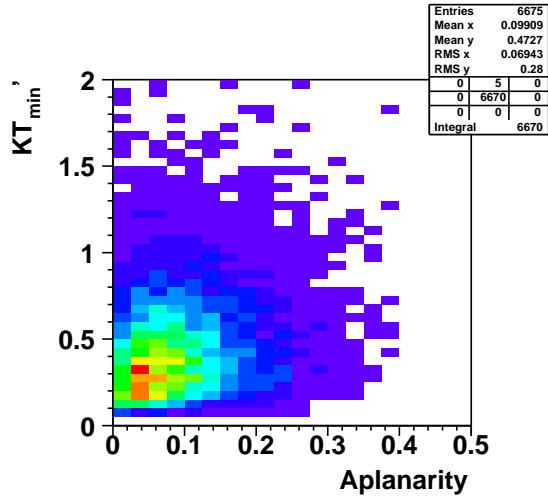Figure 23: Correlation between $\mathcal{A}$ and H'_{T2} .

Entries	6675
Mean x	0.09909
Mean y	0.4727
RMS x	0.06943
RMS y	0.28

0	5	0
0	6670	0
0	0	0
Integral		6670

Figure 24: Correlation between $\mathcal{A}$ and K'_{Tmin} .

Entries	6675
Mean x	0.09921
Mean y	62.8
RMS x	0.06946
RMS y	31.58

0	17	0
0	6658	0
0	0	0
Integral		6658
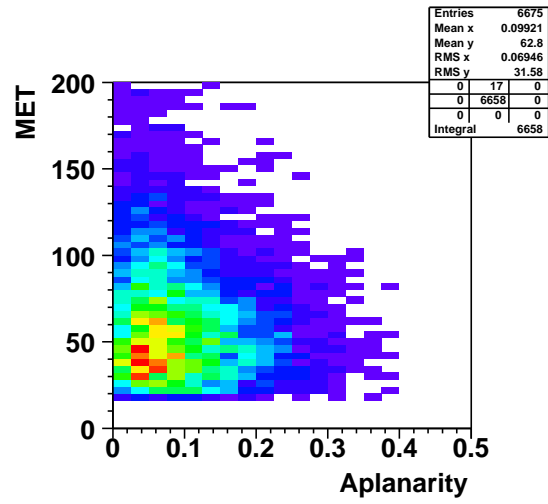
Figure 25: Correlation between $\mathcal{A}$ and $\not{E}_T$.

Entries	6675
Mean x	0.5889
Mean y	62.79
RMS x	0.2601
RMS y	31.58

0	17	0
0	6653	5
0	0	0
Integral		6653
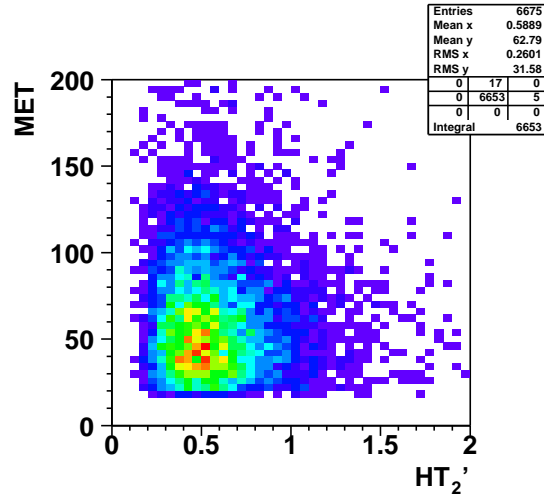
Figure 26: Correlation between H'_{T2} and $\rlap{/}{E}_T$.

Entries	6675
Mean x	0.4725
Mean y	0.5888
RMS x	0.2799
RMS y	0.2599

0	5	0
0	6665	5
0	0	0
Integral		6665
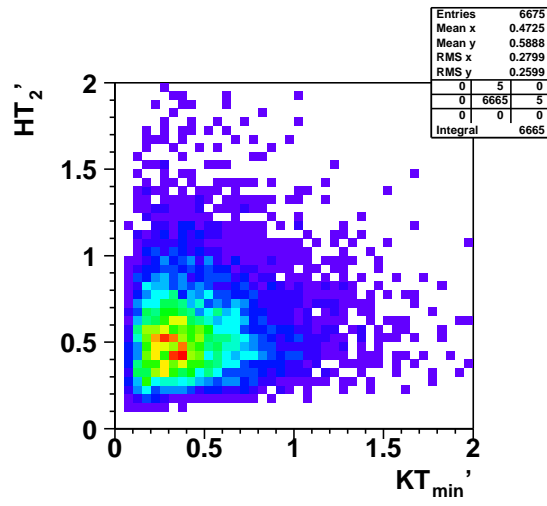
Figure 27: Correlation between K'_{Tmin} and H'_{T2} .

MET

Entries	6675
Mean x	0.4733
Mean y	62.82
RMS x	0.28
RMS y	31.58

0	17	0
0	6653	5
0	0	0
Integral	6653	

KT_{min}'
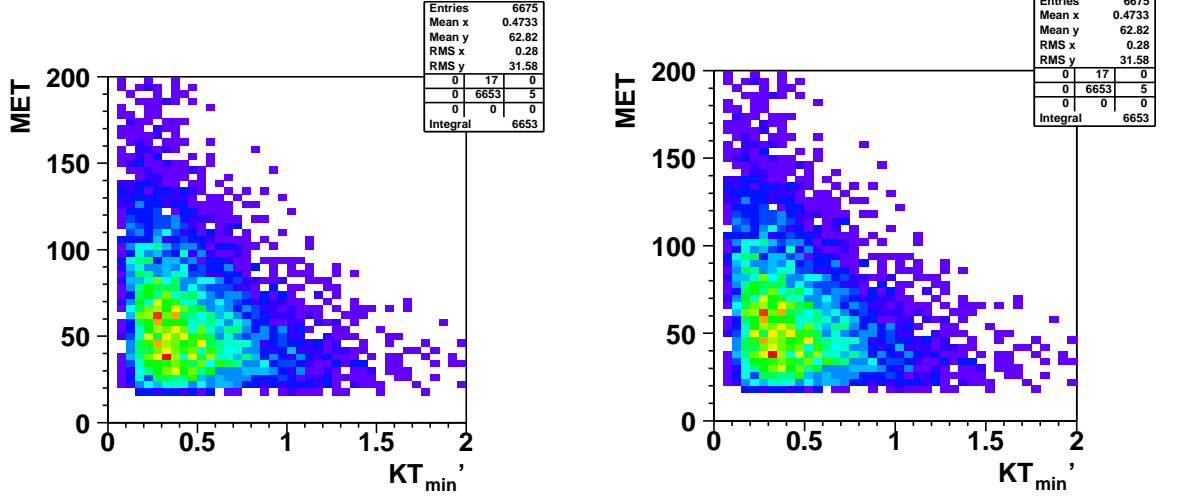
MET

KT_{min}'

Figure 28: Correlation between K'_{Tmin} and $\not{E}_T$.

3. K'_{Tmin}: $\sqrt{K'_{Tmin}}$

4. $\not{E}_T$: $e^{-\max\left(0, \sqrt{\frac{3(\not{E}_T - 5)}{2}}\right)}$

From this set of functions, then we form the bin by bin ratio for variable i :

$$L_i^{tb} = \ln\left(\frac{top}{background}\right) \tag{8}$$

The next step is to fit polynomial functions to each of the points generated above. One should also note that underflow and overflows are placed in the edge bins of the histogram. This is to ensure that there is a minimal loss of information in this procedure. The results of fit are shown in figures r-t.

From these functions a combined log log likelihood is found by summing the individual log likelihoods:

$$L^{tb} = \sum_{i=1}^{4} e_i L_i^{tb}(f_i) \tag{9}$$

Here f_i is the function of the topological variable and e_i are linear coefficients. The e_i are introduced because the variables are correlated.
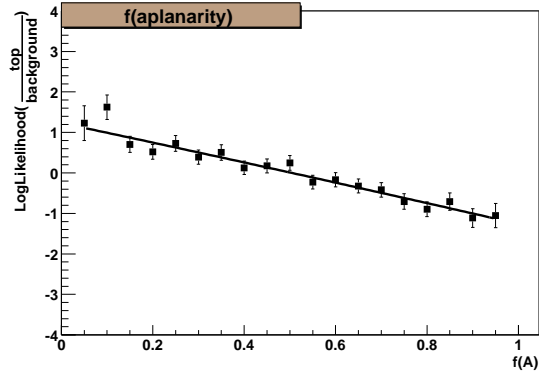
Figure 29: Quadratic fit of $\ln(\frac{top}{background})$ for $\mathcal{A}$.
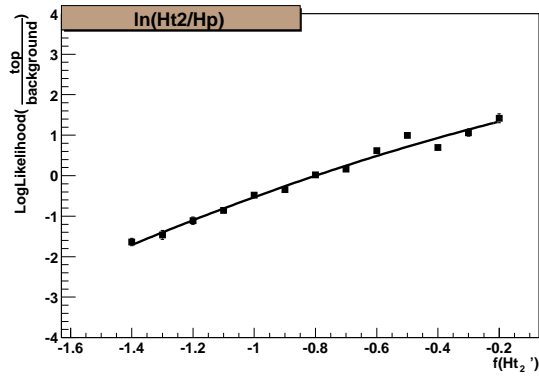

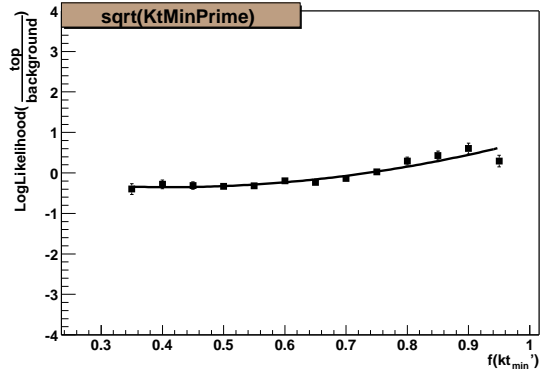Figure 30: Quadratic fit of $\ln(\frac{top}{background})$ for H'_{T2}.

Figure 31: Quadratic fit of $\ln(\frac{top}{background})$ for K'_{Tmin}.

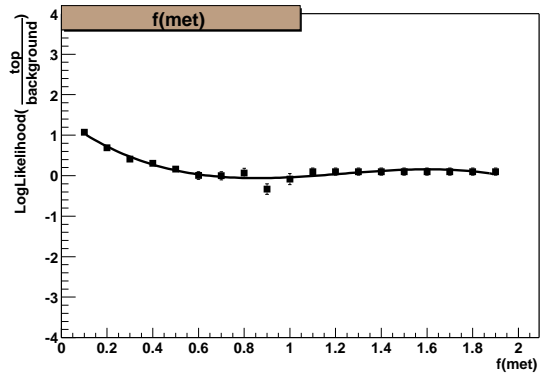Figure 32: Cubic fit of $\ln(\frac{top}{background})$ for $\displaystyle{\not}E_T$.

Methods described in [5] were developed to optimize the combination of likelihoods. For this study, we assume that the variables are sufficiently uncorrelated to produce good discrimination. Eventually, we will of course want to optimize these. At the moment, we take these coefficients to be unity.

Finally, the likelihood discriminant is formed:

$$\mathcal{D}_{LB} = \frac{1}{1 + \mathcal{P}e^{-L^{tb}}} \tag{10}$$

We choose P to be 1.25. Clearly the choice of this variable is independent of the discrimination achieved. For the moment, we choose this value to be consistent with [5].

To evaluate the discriminant, we compute the individual likelihoods for each variable by using the fit function on an event by event basis. For each event we then use equation 10 to form the final discriminant. Note that because of limited Monte Carlo at this time, we were forced to use the same sample to find the various fit parameters and to evaluate there performance. This probably accounts for the much improved discrimination that we see as compared to [5].

Since all of these correlations are below 0.1, we find these variables acceptably uncorrelated with the fit mass. In order to a quantitative measure of the separation achieved we compute the signal to noise ratios s_i for each variable i as:

$$s_i \equiv \frac{\frac{1}{2}\langle L_i^{tb}\rangle_{top} - \frac{1}{2}\langle L_i^{tb}\rangle_{bkgnd}}{\sqrt{\frac{1}{2}\sigma^2(L_i^{tb})_{top} + \frac{1}{2}\sigma^2(L_i^{tb})_{bkgnd}}} \tag{11}$$

for this we get:

1. $s(\not{E}_T) = 0.05$
2. $s(\mathcal{A}) = 0.63$
3. $s(H'_{T2}) = 0.32$
4. $s(K'_{Tmin}) = 0.28$

To compute the s_{tot} we add the above in quadrature. We obtain $s_{tot} = 0.762$.

4   Kinematic Fitting

The Standard Model predicts that the top quark decays into a bottom quark and a W boson before the top has time to hadronize and form
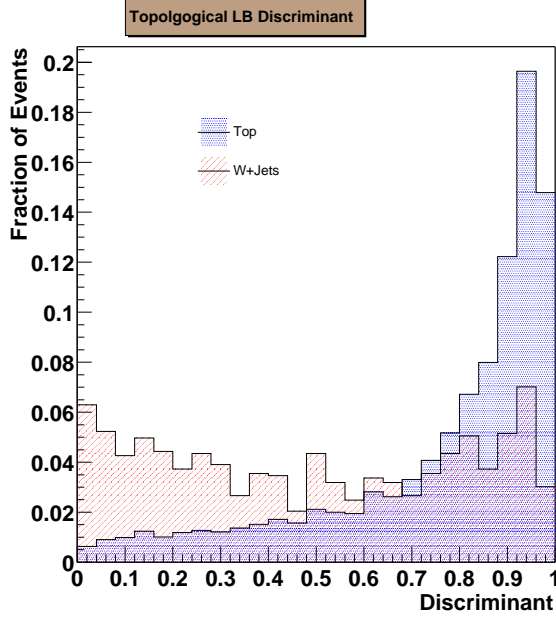
Figure 33: Discriminant for background and $t\bar{t}$.

a hadron with a branching fraction of approximately 100% . In semi-leptonic decay of the $t\bar{t}$ system, one W decays hadronically producing two light quarks while the other decays leptonically into a charged lepton and a neutrino, see figure 1. In order to reconstruct the mass of the top quark from its decay products, we need to measure the four-vectors of the final state particles: the jets, charged lepton, and neutrino. While the jets formed after the hadronization of the final state quarks and the charged lepton can be directly observed and their properties measured by the DØ detector the neutrino cannot. However, the presence of the neutrino can be inferred from the missing transverse energy in the event: $\not{E}_T$. This missing energy is taken to be the magnitude of the transverse momentum of the neutrino. However, this procedure cannot be extended to the longitudinal or z component of the neutrino momentum due to particles with very small transverse momentum that escape down the beam pipe. The p_z of the neutrino is the one kinematic unknown in the event.

There are three constraints on the four-vectors in the decay. Since the two W's are produced on shell in $t\bar{t}$ events, two jets should have invariant mass of the W as should the charged lepton and neutrino. As

31

well, the CPT theorem demands that the mass of the top and anti-top quark are the same (we make the approximation that We have three constraints but only one unknown, hence the kinematic algebra is twice over constrained. In order to start the fitting procedure, one must have an initial value for all of the kinematic variables. The p_z of the neutrino is found by forcing the invariant mass of the charged lepton and the neutrino to be that of the W boson. This leads to a quadratic equation and two possible solutions. This is consistent with over-constrained nature of the problem. In principle, one could attempt to use the results of the fits from both solutions. In practice it has been found that taking the smallest absolute value of the neutrino solution yields the correct solution in approximately 60% of the cases [10]. In the same work it was also found that either solution ultimately yields the same fit mass in over 90% of the cases.

The fit is based on the SQUAW [12] algorithm and was implemented and used extensively at DØ to analyze data from Run I. As well S. Snyder, implemented the algorithm in the DØ Run II computing framework as the hitfit package. The fit is performed by minimizing a χ^2 defined as:

$$\chi^2 = (\vec{x} - \vec{x}_M) G (\vec{x} - \vec{x}_M)^T \tag{12}$$

where $\vec{x}_M$ is a vector of measured variables, $\vec{x}$ is a vector of fitted variables, and G is the inverse error matrix of the measured quantities. G is taken to be diagonal. The fit attempts to minimize the χ^2 subject to the three constraints.

If the fit were linear in its parameters, one could solve the minimization problem by the method of Lagrange Multipliers and standard linear algebra techniques. Since this is not the case an iterative technique is used. The starting point is the vector of measured values of the variables. Then the constraint functions are expanded in a power series and linearized. Using Lagrange Multipliers the χ^2 is minimized with respect to linearized constraint equations. The fit values for that iteration are taken as the starting point for the next iteration. In the next iteration the constraint equations are re-evaluated. This process continues until the constraint equations are satisfied or the χ^2 stops changing. If the number of iterations is larger than a max-iterations value parameter without converging that fit is discarded. For a more detailed description of the algorithm see [10].

Another fold in the procedure is jet permutations. To lowest order, there are four jets from the hadronization of the four quarks in the final state. A priori it is not known which jet came from which parton.

Therefore, one has to attempt all the possible jet assignments. With four jets there are twelve distinct permutations (in principle there are twenty-four permutations for four objects but the invariant mass of the hadronically decaying W is the same if the two jets assigned to the W are exchanged). These incorrect permutations are another source of background. Often initial and final state gluon radiation can lead to more than the nominal four jets. Currently, as was the case in Run I, only the four highest p_T jets are used in the fit. Currently, we assume that any other jet is from initial state radiation.

The most important aspect in the mass measurement is extracting the top mass from a data sample which contains both top signal and background. To do this, we use the technique developed in [4]. This is done using a binned Poisson-statistics maximum likelihood fit. Since we cannot generate different Monte Carlo samples with a continuous input mass, we are forced to use a set of discrete top quark masses and interpolate. This procedure was developed and described in detail in [7]. Here we only give a brief introduction and focus on our analysis.

From the preselected sample, defined previously we attempt to get further discrimination between top quark events and background by forming a discriminant. The range of the discriminant is from 0 to 1: the closer to one the more top like, the closer to zero the more the event resembles background. In the present analysis we use the discriminant to cut away background by only using events that pass the discriminant criterion.

The fit that is performed is likelihood fit which seeks to estimate both the true top quark mass and the fraction of top in our data sample. From the w+jet Monte Carlo we obtain a histogram represents what we would observe without the presence of a top quark signal in the data. As well, using Monte Carlo we generate a series of signal templates at various input masses ranging from 150 to 200 GeV. The purpose of the likelihood fit is to give our best estimate of which template gives the best representation of our data.

The likelihood function is a product of Poisson factors.

$$L = L(m_t, A_s, A_b) \times L(N, \bar{N}) \qquad (13)$$

where:

$$L(m_t, A_s, A_b) = N \frac{e^{-\bar{n}}}{n!} \qquad (14)$$

Here, N is the number of observed events and n is defined as:

$$n = A_s n_s + A_b n_b \qquad (15)$$

here A_s and A_b are the signal and background strength respectively, while ns and nb are the number of signal events and number of background events that are being simultaneously fit for along with the top mass.

The second term in equation one, is a Poisson term which constrains the number of background events to the expected number. It can be shown that this can be expressed as:

$$\bar{N} = n_b(1 + \frac{ns}{nb}) \qquad (16)$$

We began by using the somewhat more sophisticated Bayesian likelihood that is described in [4] and [7]. However, we soon found that there was little practical advantage to that approach and because of the increased computationaltime needed to minimise the much more complicated form of the likelihood we chose for the simpler likelihood described here.

5 Likelihood Fitting Procedure

Each event has a fit mass and a LB discriminant. If the event passes the discriminant cut, the fit mass (obtained from the hitfit kinematic fitting package) is binned from 80 to 280 GeV. Because of the low statistics involved, we use 20 bins with an even width of 10 GeV. The same procedure is done with each sample of top Monte Carlo samples. These make up our signal templates representing different possible top masses. After the full event selection on the dedicated multi-jet sample, only 65 events in the μ + jets sample and 40 in the electron channel survive, before any selection based on the low bias likelihood. Therefore, this template histogram is only formed from the W+jets Monte Carlo.

We have a series of 7 template mass points whose values are 150, 160, 170, 175, 180, 190, and 200 GeV. Because of the limited statistics, we decided to combined both the electron and muon channel into one analysis. Therefore, we mixed the templates for both signal and background in the appropriate ratio.

As mentioned before, the likelihood fit is done using a binned Poisson method. For each mass point, we use the number of entries in each

mass bin in the template histogram and the background histogram (properly normalized) to estimate the likelihood of obtaining the observed number of entries given our expected background in that bin and the hypothesized contribution from a top mass at that generated value. The free parameters of the fit at each mass point are the fitted number of signal and background events. The actually fit is carried out by the Minuit minimization package. It finds the value of the number of signal events and background events that maximizes the likelihood (or equivalently minimizes the negative log likelihood). As well it returns the minimum likelihood.

Since we have a finite amount of statistics for each template point, the value of the likelihood at each mass template has an error associated with it. Because of the limited Monte Carlo statistics, some of the mass points show significant statistical fluctuation. After the minimum likelihood has been found, we scale each bin up and down by the square root of the number of entries in that bin. The difference in the observed likelihood from this procedure is recorded and the sameprocedure is done for each bin. The resulting variation is added in quadrature and is quoted as the error on the likelihood.

The same procedure is repeated for all template points. The minimum value of the negative log likelihood, its error, and the fitted number of signal and background events are recorded. These points are then used to find the global minimum in the the negative log likelihood. A parabola is then fit to the resulting points to extract the most likely top mass, its error (where the - log likelihood rises by 0.5 from its minimum) and the fit number of signal and background events.

6 Ensemble Testing

Given the number of top events is the data set is quite small, it is important to investigate carefully the statistical properties of the top mass estimator. The ensemble testing procedure is taken from the top group's prescription to compare different mass extraction techniques. In the following ensemble tests, the number of events in the ensemble is fixed to be the observed number of events in data 137. Then the composition for each ensemble is taken from the estimated composition of the data sample as shown in the event yield section. In each ensemble formed the signal and background events are selected before the low bias likelihood selection is used. Then the events which pass
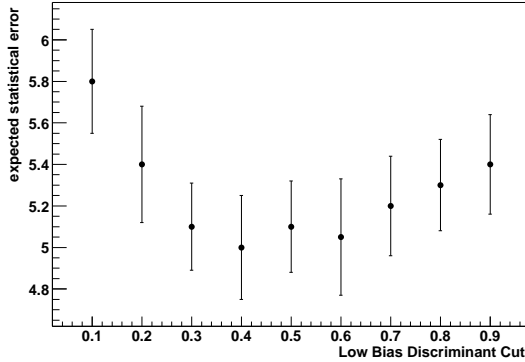
Figure 34: Expected error for various requirements on the low bias discriminant.

the low bias likelihood selection are used for the fit in each ensemble in theidentical way that data is treated. Because of the limited number of multi-jet events discussed in the previous section, and the very small contribution to the total background the multi-jet events were not used in either the background template nor in the ensembles (w+jet events were used in place). In order to determine the optimum selection criterion for the low bias discriminant, the requirement on the low bias discriminant was varied and the mean value of the expected statistical error was computed. The result is shown in figure 34 indicating that the optimum selection was around 0.4.

Since the likelihood method relies on templates, one expects that the method should not need to calibrate the likelihood. For each input mass, used to create the ensembles, we should be able to get a result that (within error) is consistent with the input. Figure 35 shows the result of the calibration curve. As expected it is consent with a slope of 1 and offset of 0.

The next check is that the method for extracting the mass gives an error that can be interpreted as a 1.0 σ confidence level interval. In order to check this, the pull distribution should have a width of 1.0 indicating that the error can be correctly interpreted. The pull distribution for a typical ensemble experiment is shown in figure 36.

Finally, the ensemble test can also be used to get an estimate of the expected statistical error. This is done by forming many ensembles and looking at the statistical error that is calculated from the likeli-
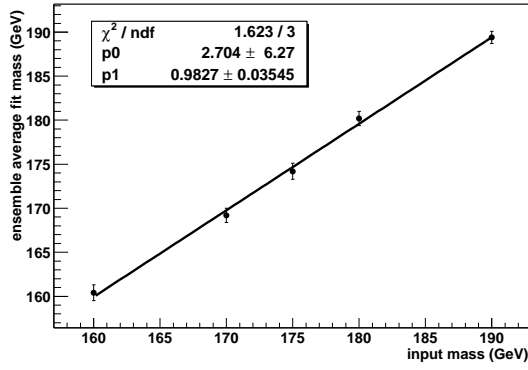
Figure 35: Calibration curve for the top mass likelihood fitting procedure.
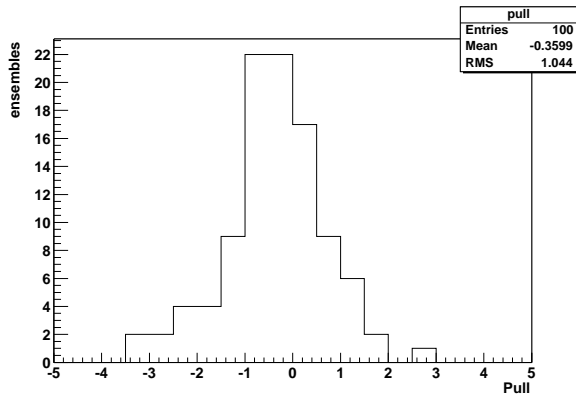

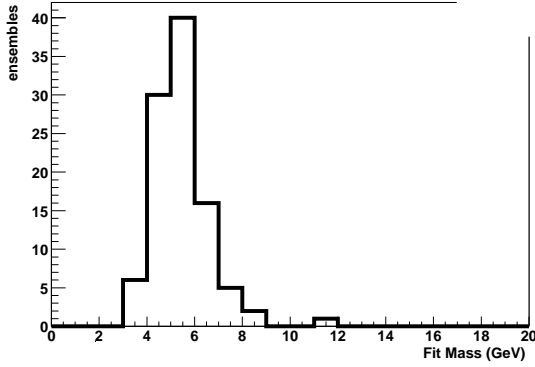Figure 36: Pull distribution for ensemble test with a rms consistent with 1.

Figure 37: Expected statistical error from ensemble tests.

hood procedure. Since the pull distribution is close to 1 this ensures that the error assigned to each ensemble is consistent with the statistical spread of the extracted top mass. Figure 37 shows a distribution of the expected statistical error.

7 Results on Data

There were 72 events out of 137 that pass the requirement of having a low bias discriminant > 0.4. The fit mass of all the preselected candidate events is shown in figure 39.

The distribution of the low bias discriminant can be seen in figure 38. The various contributions are from the expected number of top, W +jet, and multijet events in the expected fraction.

In figure 40 the distribution of the fit mass events that pass the low bias discriminant selection are shown. As described earlier, the fit also returns an estimate of the most likely value of the signal and background events. The corresponding fit contribution is also plotted and normalized to the data. For the top mass contribution, the template closest to the fit mass extracted from data is plotted.

The results from the likelihood curve give a top mass of 168 GeV $\pm$ 13 GeV (statistical).

In the final figure, the fit mass distribution for those events which fail the low bias cut is shown. Note that in both the signal and background region (LB > 0.4 and LB < 0.4) the model for the data fails to predict the amount of observed low mass events.
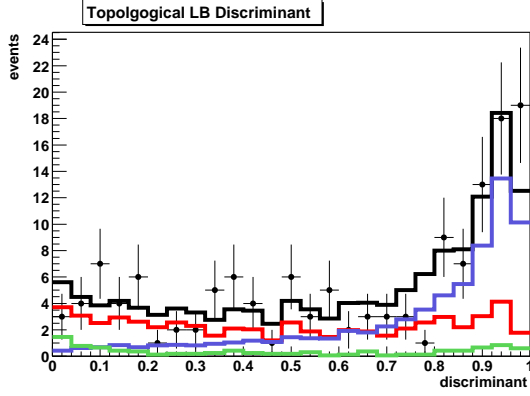
38

Figure 38: Discriminant for the 137 selected data events. In red is the expected contribution from W + jets, in blue is the expected contribution from $t\bar{t}$ and in green is the expected qcd contribution.
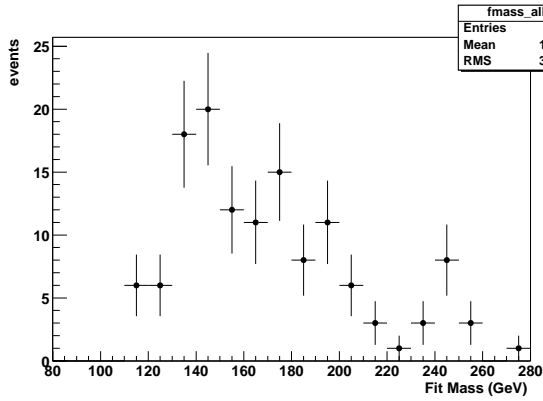

Figure 39: Fit mass distribution in the lowest χ^2 combination for all events passing the event preselection.
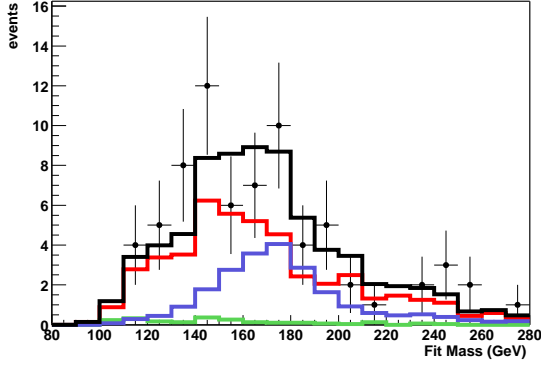
Figure 40: Fit mass distribution in the lowest χ^2 for the events that pass the low bias discriminant requirement. As well, the contributions from signal and background are plotted in the fraction returned from the fit on data. In blue is the contribution from $t\bar{t}$ events while in red and green are the contributions from W+jets and multi-jet events respectively. In black is the expected contribution from all sources normalized to data.
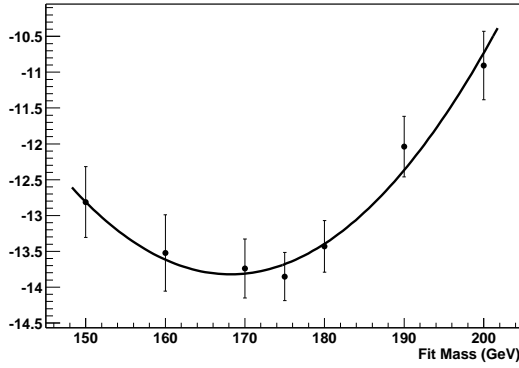

Figure 41: Likelihood curve fit to the lepton + jets data used to estimate the most likely value of the top mass and its error.
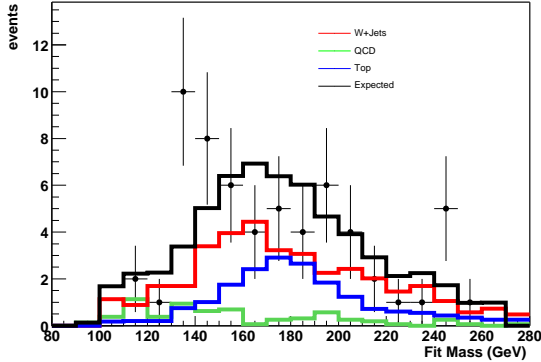
Figure 42: Fit mass distribution for the events with LB < 0.4. In red is the expected contribution from W + jets, in blue is the expected contribution from $t\bar{t}$ and in green is the expected qcd contribution.

7.1 Cross Checks

In order to cross check the results of the fit to seperate event selections were tried. The first was to require the same event selection, but to also require that at least one jet had an secondary vertex (SVX) assoicated with the jet. The algorithm is described in [13]. Here we make the same requirement that is used in the cross-section with btagging, using the tight definition. This is expected to greatly increase the signal to background contribution. The fit mass of the selected events passing the LB discriminat selection of 0.4 is shown in figure 43.

As well, the discriminant distribution for the tagged events supports the there is indeed a high fraction of top in the sample.

In this analysis, we have not considered btagging into the fit information. In order to do so, we would need to do detailed studies to study the properties of the tagged sample, the signal composition, and the expected resolution of the using the tagging information into the fit. Here, we have a much more modest goal and simply use this information to further supress background and do not use the b quark identification information into the fit. This gives us a purer sample to cross-check that the background is understood and removing it gives us a consistent picture. The results are shown in figure 41.

One further cross-check was made to study the stability of the fit. Since all methods start with the same preselection but then apply
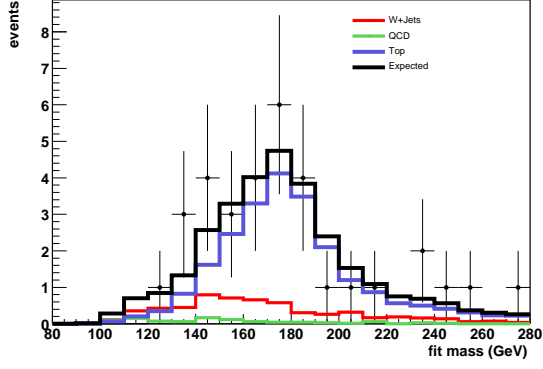
Figure 43: Fit mass distribution of the events which pass all other event requirements and have at least one jet that is identified as a bjet.
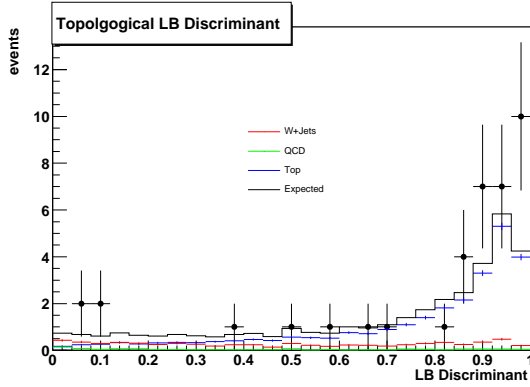

Figure 44: Low bias discriminant for the tagged events. Note the obviously large excess in the top like region as expected.
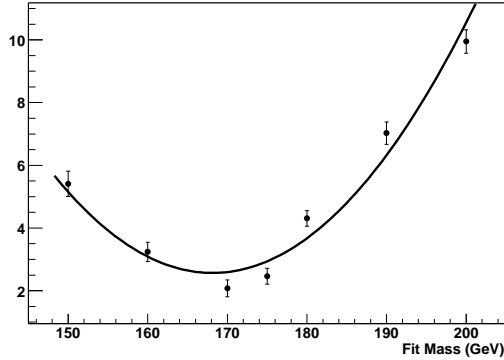
Figure 45: The result of the likelihood fit of the tagged subset of the data. The likelihood curve gives a fit mass of 169 GeV ± 9 GeV, in remarkable agreement with the untagged sample.

different methods for obtaining a higher signal to background, we also looked at what result the kinematic fitting procedure would yeild on the events that they selected. By chosing the events that passed the background probability cut and had exactly four jets we obtain a consistent result. Figure 46 shows the fit mass distribution of the matrix element selected events which also have a solution for the top mass according to the kinematic fitting. The resulting fit yeilds a top mass of 178 ± 9 GeV. Note that there is overlap between the two samples however, since our sample includes events with more then four jets.

As well the matrix element background probability cut yeilds a spectrum of events which have high discriminat values indicating that that both methods agree on which events are more top like.

These two cross-checks are important to establish that if the event selection had been different the result, within error, would have been consistent. Seeing that they are completely consistent within error increases the confidence in the fit results.

7.2   Systematic Error

The most important systematic error for the top mass in the lepton + jets channel is that associated with the jet energy scale. The error associated with this is estimated in the following procedure. Monte
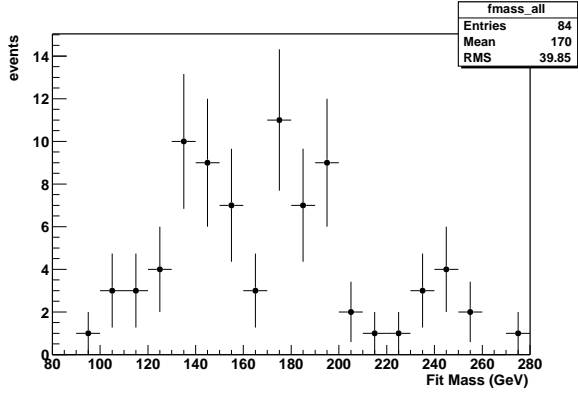
Figure 46: The fit mass distritribution of events of the events selected by the matrix element
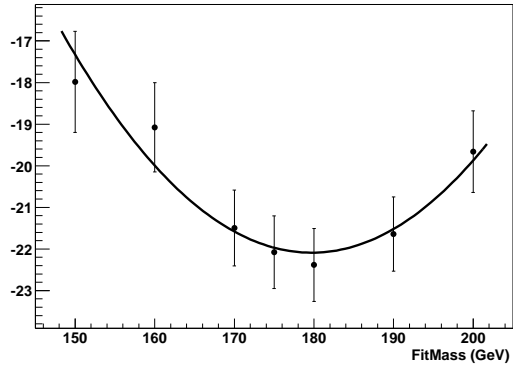

Figure 47: The likelihood curve of the events selected by the matrix element background probability cut and fit with the kinematic fitting routine.
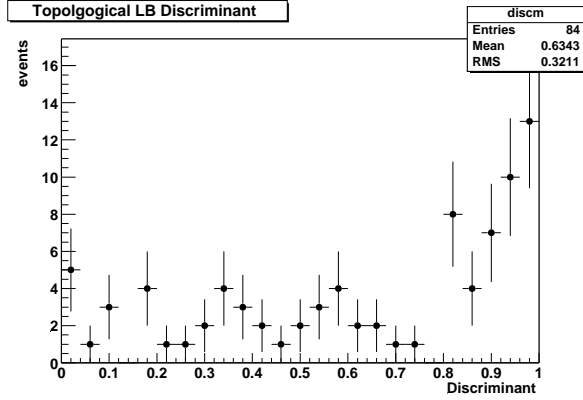
Figure 48: The distribution of the LB discriminant for those events which pass the matrix element background probability calculation and have exactly four jets.

Source	-1σ variation	$+1\sigma$ variation	average
Jet Energy Scale	4.9 GeV	7.4 GeV	6.2 GeV
Jet Energy Resolution	1.4 GeV	3.2 GeV	2.3 GeV
Event Modeling	2.0 GeV	4.0 GeV	3.0 GeV
Underlying Event (TuneA)	–	–	3.0 GeV
Limited MC Statistics	–	–	0.5 GeV
Trigger Bias	–	–	1.0 GeV
Calibration	–	–	0.5 GeV

Table 1: Systematic Errors

45

Carlo samples where the jet energy scale has been scaled up and down by $\pm 1 \sigma$ have been produced. These samples were then used to form ensembles while the templates with the nominal jet energy scale were used for the templates. This procedure yeilds an average varaition of 6.2 GeV.

Another systematic error that was considered was the variation due to the underlying event. Unfortunately, all the Monte Carlo events for the top group were generated with multiple parton interactions turned off. In order to estimate the size of the effect on the top mass samples for the central mass point of 175 GeV and the W+jets were generated with the multiple parton interactions turned on. The same procedure for the jet energy scale systematic was followed and a 3.0 GeV shift in the average value of the mass was observed from ensemble tests.

One of the largest concerns about the data in RunII is the quality and resolution of the jets. In order to to see the effects of the jet energy resolution in data, monte carlo was prepared with the jet energy smearing changed by a factor of $\pm 1 \sigma$. The variation found in using building ensembles from these energies was found to be 2.3 GeV.

In order to understand how the event modeling effects the expected error, we changed the background composition by $\pm 50 \%$. The resulting variation on the statisistical error from the ensembles was 3.0 GeV.

Although the calibration curve is consitent with with zero offset and unit slope, there is of course point to point variation in the calibration curve. We take the parameters and from the fit calibration along with the errors and compute the error which is associated with the uncertainty in the calibration curve. This procedure produces an uncertainty of 0.5 GeV

Since there is a finite amount of Monte Carlo statitistics, one expects that the statistical fluctuations in the signal and background templates could lead to variations in the extracted top mass. In order to quantify this effect we divided the Monte Carlo into four subsamples to produce four different sets of background and signal templates. Then from the full Monte Carlo set, ensembles were produced and fit with the various sets of signal and background templates. Note that the ensembles formed were fixed and the templates sets were varied. The resulting varation was then noted and divided by the square root of the number of different template $\sqrt{N-1}$ (e.g N=4 here). The result is 0.5 GeV.

Finally, in order to study the trigger efficiency bias, we will vary the

various components of the trigger by $\pm 1 \sigma$ and do ensemble tests to see what effect each has on the average ensemble mass. For the moment, we simply prepare monte carlo ensembles where the trigger efficiency is not used to wieght the monte carlo events. To be conservative, we take there error to be twice the variation that is obtained and quote an error of 1.0 GeV untill more refined studies can be done.

Adding all of these in quadrature we find a systematic error of 8.0 GeV.

References

[1] DØ Top Analysis and Data Sample for the Winter Conferences 2004, The Top Phyiscs Working Group of the DØ collaboration, DØ Note xxxx, Febuary 2004

[2] The electron likelihood in p14, J. Kozminiski et all, November 2004

[3] Measurement of the $t\bar{t}$ Production Cross-Section at $sqrts = 1.96$ TeV in the Lepton + jets Final States usinag a Topological Method, Top Production Group, Feb 2004

[4] *Direct Measurement of the Top Quark Mass by the DØ Collaboration*, PDF, Phys. Rev. D **58**, 052001 (1998)

[5] Details of Run 1 "top likelihood" estimation for top quark mass analysis in the lepton+jets channel, Mark Strovink, DØ note 3989, May 2002

[6] Collider Physics, V. Barger and R. Phillips, Addison-Wesley, NY, 1987

[7] *Bayesian analysis of multisource data*, P.C Bhat, H.B. Prosper, and S.Snyder, Phys. Lett. B **407**, 73 (1997)

[8] S. Baker and R. Cousins, Nucl. Instrum. Methods Phys Res. , **221**, 437 (1984)

[9] R. Barlow, SLUO Lectures on Statistics and Numerical Methods in HEP, Resampling and the Bootstrap, Lecture 6, (2000)

[10] *Measurement of the Top Quark mass at DØ*, Scott Snyder, Doctoral Thesis, State University at Stony Brook, NY, May 1995

[11] Particle Data Group W properties, K. Hagiwara et al., Phys. Rev. **D66**, 010001 (2002)

[12] J. Peter Berge, Frank T. Solmitz, and Horace D. Taft. *Review of Scientific Instruments*

[13] Ariel Swartzman, M. Narain, b-quark jet identification via Secondary Vertex Reconstruction, DØ note 4080, Jan 2003